

Causal Inference (MGT-416)

Final Project

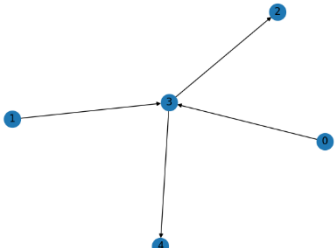
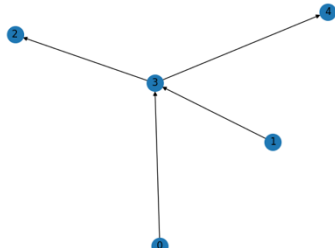
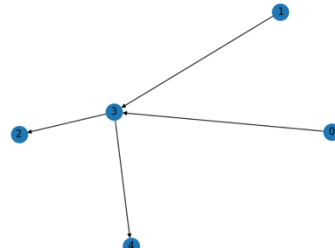
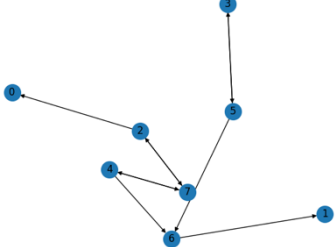
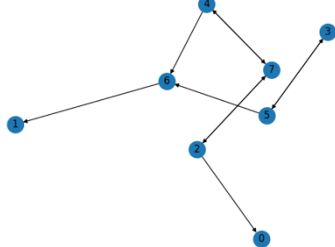
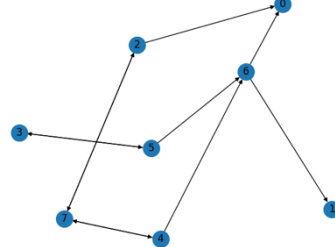
1. Learning the causal structure of a Bayesian network

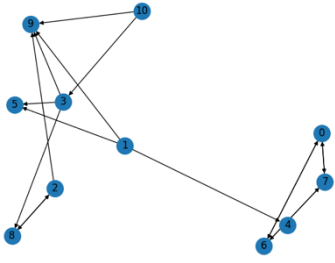
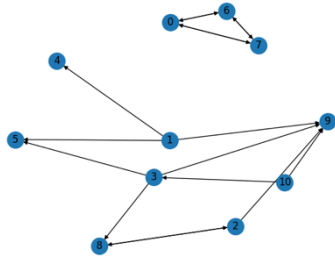
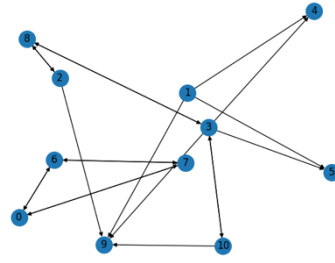
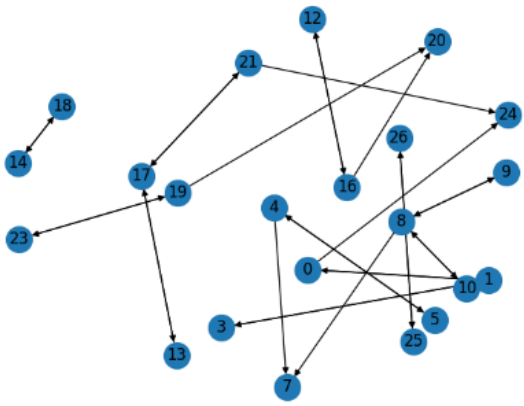
Task 3 & Task 4

As required in task 4 we calculated the computational complexity (i.e., maximum number of CI tests each algorithm performs in the worst case) for the three functions SGS, PC1 and PC2:

| SGS: | PC1: | PC2: |
|---------------------|---------------------|-------------------------------|
| $C_2^n * 2^{(n-2)}$ | $C_2^n * 2^{(n-2)}$ | $C_2^n + (C_2^n * 2^{(n-2)})$ |

The following table summarizes the results obtained after performing task 3 & 4:

| Dataset #1 | | |
|--|--|--|
| Algorithm: SGS | Algorithm: PC1 | Algorithm: PC2 |
| Number of performed tests: 51 | Number of performed tests: 51 | Number of performed tests: 42 |
| Computational complexity: 80 | Computational complexity: 80 | Computational complexity: 90 |
| Max. degree of the graph: 4 | Max. degree of the graph: 4 | Max. degree of the graph: 4 |
| Oriented graph:  | Oriented graph:  | Oriented graph:  |
| Dataset #2 | | |
| Algorithm: SGS | Algorithm: PC1 | Algorithm: PC2 |
| Number of performed tests: 607 | Number of performed tests: 551 | Number of performed tests: 483 |
| Computational complexity: 1.792 | Computational complexity: 1.792 | Computational complexity: 1.820 |
| Max. degree of the graph: 3 | Max. degree of the graph: 3 | Max. degree of the graph: 4 |
| Oriented graph:  | Oriented graph:  | Oriented graph:  |

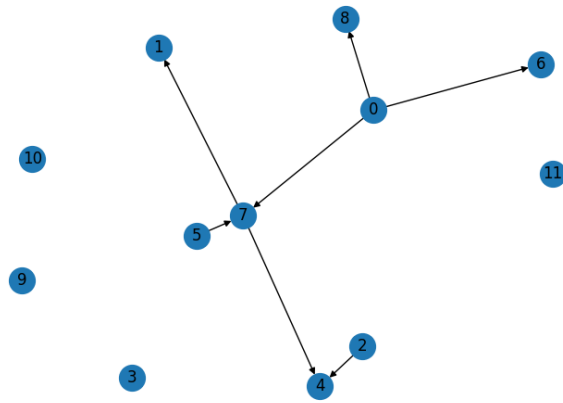
| Dataset #3 | | |
|--|---|---|
| Algorithm: SGS | Algorithm: PC1 | Algorithm: PC2 |
| Number of performed tests: 7.501 | Number of performed tests: 5.681 | Number of performed tests: 5.380 |
| Computational complexity: 28.160 | Computational complexity: 28.160 | Computational complexity: 28.215 |
| Max. degree of the graph: 4 | Max. degree of the graph: 4 | Max. degree of the graph: 5 |
|  |  |  |
| Dataset #4 | | |
| Oriented graph: | | Algorithm: PC2 |
|  | | Number of performed tests: 1.801.091 |
| | | Computational complexity: 11.777.605.983 |
| | | Max. degree of the graph: 3 |
| | | |

- Note that undirected connections are shown as bidirectional arrows
- The partially-oriented matrix for PC2-D4 is attached with the name "D4.csv"

We can notice that the worst-case complexity is never reached. Indeed, in our simulations the number of CI tests performed is always lower compared to the complexity of the algorithm. Hence, in the provided causal patterns, there is never a complete connection between all the nodes.

2. Analyzing the stock market data

Learned directed information graph



Legend

| Company Name | Node number |
|----------------------|-------------|
| Apple Inc. | 0 |
| Cisco Systems Inc. | 1 |
| Dell Inc. Inc. | 2 |
| EMC Corporation | 3 |
| Google Inc. | 4 |
| Hewlett-Packard | 5 |
| Intel | 6 |
| Microsoft | 7 |
| Oracle International | 8 |
| Business Machines | 9 |
| Texas Instruments | 10 |
| Xerox | 11 |

Directed Information matrix (12 × 12)

DI Matrix:

| | | | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.000 | 0.282 | 0.250 | 0.290 | 0.466 | 0.499 | 0.637 | 0.678 | 0.552 | 0.383 | 0.380 | 0.175 |
| 0.337 | 0.000 | 0.251 | 0.428 | 0.319 | 0.255 | 0.087 | 0.291 | 0.243 | 0.191 | 0.174 | 0.159 |
| 0.212 | 0.384 | 0.000 | 0.161 | 0.571 | 0.143 | 0.234 | 0.400 | 0.249 | 0.436 | 0.345 | 0.276 |
| 0.200 | 0.181 | 0.318 | 0.000 | 0.277 | 0.159 | 0.238 | 0.309 | 0.420 | 0.252 | 0.207 | 0.277 |
| 0.344 | 0.274 | 0.266 | 0.336 | 0.000 | 0.265 | 0.317 | 0.299 | 0.337 | 0.399 | 0.365 | 0.247 |
| 0.350 | 0.415 | 0.200 | 0.340 | 0.300 | 0.000 | 0.197 | 0.635 | 0.224 | 0.425 | 0.248 | 0.327 |
| 0.148 | 0.337 | 0.313 | 0.323 | 0.132 | 0.298 | 0.000 | 0.179 | 0.198 | 0.214 | 0.350 | 0.245 |
| 0.480 | 0.548 | 0.340 | 0.360 | 0.542 | 0.210 | 0.368 | 0.000 | 0.205 | 0.284 | 0.382 | 0.408 |
| 0.436 | 0.342 | 0.250 | 0.100 | 0.319 | 0.131 | 0.318 | 0.199 | 0.000 | 0.268 | 0.273 | 0.257 |
| 0.290 | 0.412 | 0.473 | 0.477 | 0.303 | 0.270 | 0.217 | 0.423 | 0.319 | 0.000 | 0.184 | 0.212 |
| 0.230 | 0.260 | 0.152 | 0.164 | 0.244 | 0.242 | 0.136 | 0.121 | 0.087 | 0.169 | 0.000 | 0.184 |
| 0.111 | 0.145 | 0.230 | 0.240 | 0.107 | 0.166 | 0.171 | 0.155 | 0.237 | 0.246 | 0.292 | 0.000 |

This matrix is attached with the file name "DI_matrix.csv"

Matrix for oriented graph:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

This matrix is attached with the file name "DI_matrix_for_graph.csv"